

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2002, 05:49:49 ; Search time 27.25 Seconds

(without alignments)
446.763 Million cell updates/sec

Title: US-09-652-292-2

Perfect score: 2765

Sequence: 1 MGHSPVPLCLASVSLGL.....GHRNSTGIPYRIEISNAS 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTOTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	438	15.8	584	2	US-08-928-692-13
2	427	15.4	493	2	US-09-031-392-10
3	427	15.4	493	4	US-09-299-549-10
4	407.5	14.7	509	2	US-09-031-392-6
5	407.5	14.7	509	4	US-09-299-549-6
6	404.5	14.6	494	2	US-09-031-392-5
7	404.5	14.6	494	4	US-09-299-549-5
8	394.5	14.3	524	2	US-08-928-692-12
9	389.5	14.1	534	2	US-09-031-392-4
10	389.5	14.1	534	4	US-09-299-549-4
11	364.5	13.2	492	2	US-08-355-844-3
12	364.5	13.2	492	5	PCT-US95-16126-3
13	349.5	12.6	500	2	US-09-031-392-7
14	349.5	12.6	500	4	US-09-299-549-7
15	331	12.0	488	2	US-08-928-692-10
16	311	11.2	383	2	US-09-031-392-3
17	311	11.2	383	4	US-09-299-549-3
18	284.5	10.3	563	2	US-09-031-392-2
19	284.5	10.3	563	4	US-09-299-549-2
20	268.5	9.7	488	2	US-08-928-692-11
21	225.5	8.2	553	2	US-08-501-572-3
22	225.5	8.2	553	3	US-09-040-444-3
23	198.5	7.2	553	3	US-08-501-572-2
24	198.5	7.2	553	3	US-09-040-444-2
25	188	6.8	556	3	US-08-501-572-1
26	188	6.8	556	3	US-09-040-444-1
27	144	5.2	537	2	US-08-647-397-2

28	140.5	5.1	423	2	US-08-494-907-14	Sequence 14, Appl
29	140.5	5.1	423	5	PCT-US96-10986-14	Sequence 14, Appl
30	136	4.9	553	4	US-09-020-956-113	Sequence 113, App
31	136	4.9	553	4	US-09-030-607-113	Sequence 113, App
32	132.5	4.8	528	2	US-08-403-852D-21	Sequence 21, Appl
33	132.5	4.8	528	3	US-08-510-646B-22	Sequence 22, Appl
34	132.5	4.8	528	4	US-09-231-818-21	Sequence 21, Appl
35	125.5	4.5	520	4	US-08-964-137-2	Sequence 2, Appl
36	118	4.3	432	2	US-08-677-049-8	Sequence 8, Appl
37	116.5	4.2	366	2	US-08-850-860-4	Sequence 4, Appl
38	116.5	4.2	396	2	US-08-944-916-4	Sequence 4, Appl
39	116.5	4.2	396	2	US-08-814-877-4	Sequence 4, Appl
40	115	4.2	416	4	US-09-333-208-2	Sequence 2, Appl
41	115	4.2	416	4	US-09-333-254-2	Sequence 2, Appl
42	115	4.2	1093	3	US-08-545-860D-55	Sequence 55, Appl
43	115	4.2	1093	5	PCT-US94-04496-55	Sequence 55, Appl
44	114	4.1	4551	3	US-09-320-878-1	Sequence 1, Appl
45	114	4.1	4613	4	US-09-105-537-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-08-928-692-13
Sequence 13, Application US/08928692
Patent No. 5958727
GENERAL INFORMATION:
APPLICANT: Brody, Howard
APPLICANT: Laver, Deborah S.
APPLICANT: Lamsa, Michael
APPLICANT: Hansen, Kim
TITLE OF INVENTION: Methods for Modifying the Production of
TITLE OF INVENTION: a Polypeptide
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59587270 No. 5958727disk of No. 5958727th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928.692
FILING DATE: 12-SEPT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lamberts, Elias J
REGISTRATION NUMBER: 33,728
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 5958727e
US-08-928-692-13

Query Match 15.8%; Score 438; DB 2; Length 584;
Best local similarity 24.9%; Pred. No. 4.8e-31;
Matches 136; Conservative 85; Mismatches 218; Indels 108; Gaps 12;
OY 4 SPVPLCLASVSLGLTFGLAVISGALLPLQDFTGSLCT--DQFLVGSLLGALL 60

```

      83 SPFIITL-TFVASISGFMGYDMGYISSALLISICTDHDHVLTYGEEIYAAVSLGLI 141
      QY 61 ASIVGGLIDCYGRKQALIGSLVLLAGSLTIGLACSLANVLGKAVGPAISLSMACC 120
      DB 142 TSIFAGAAADIFGRKRCRSMNMFVIGAILQVSAHTFMQMGRLMGVIGSLIAP 201
      QY 121 IYSELVPGROGVLSLYENGITVIGIISLYALNVALAGTPMGCRHMGATADAVLOSL 180
      DB 202 LFIEIAPKMIKRGRLTYINSIMLTGQIVAYCGAGLNYNMGMRLIVGLSLPTAVQFT 261
      QY 181 SLFLP-----AGTDEATH-----KDLIFLOGGEAPKLGGRP 214
      DB 262 CLAFELPTPRYYMKGLARATEVILKRSYTDTSSEIERKVEELVTLNDSIPGRNPEKV 321
      QY 215 RYSEFLDFRARDNKRRTTVGLVLFQOLTGOPNVLYCASTIFFSSVGFHSSAVLASV 274
      DB 322 MNTIKELHTVPSNLRALICGLQAIQOFTGWSNLMFTSGTIFETVGR-NSSAV--SI 377
      QY 275 GLGAVKVAATLTAMGLVDRAGRALLAGC--ALMAISVSGIGLVSPAVPMDSGSCSLAV 332
      DB 378 IYSGTNEITLVAFESIDKIGRTILILGPGMTMALVYCSTIAFHFLGIKFD-GAVAVV 436
      QY 333 PNATGOTGLPDSGLDSSLPPIPTNEDQREPLISTAKTKRPHRSGDPSAPRLALS 392
      DB 437 SSGFSSMGI-----445
      QY 393 SALPGPLPARGHALLRMTALLCLMVFVSASFSGFPYTWLVASEIYPVEIRGAPFCN 452
      DB 446 -----VLIFFIVEAFAFALIGITVPM-QOSELPQNVNGIGISYAT 486
      QY 453 SFNMANLFISFLDILGTSMTFLVGLTAVLGIFLYLFEPKRGOSLAETDOOF 512
      DB 487 ATNMAGSLVIASTFLMQLNITPAGTFAFGISCLSTIFCYFCYPELSGLEEV-QYI 545
      QY 513 OKRRFTL 519
      DB 546 LKDGFINI 552
  
```

RESULT 2
 US-09-031-392-10
 Sequence 10, Application US/09031392
 Patent No. 5942398
 GENERAL INFORMATION:
 APPLICANT: Tartaglia, Louis A.
 APPLICANT: Meng, Xun
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/031.392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkiohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154

```

      ; INFORMATION FOR SEQ ID NO: 10:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 493 amino acids
      ; TYPE: amino acid
      ; TOPOLOGY: linear
      ; MOLECULE TYPE: protein
      ; US-09-031-392-10
  
```

Query Match 15.4%; Score 427; DB 2; Length 493;
 Best Local Similarity 25.9%; Pred. No. 3,76-30;
 Matches 144; Conservative 80; Mismatches 104; Indels 148; Gaps 15;

```

      QY 8 LPLCAVSLLGLTFYELAVISG-----ALPLQIDFGSLCEQ 47
      DB 14 LVLAVALAALGSPQYGNLGVINAPQVTEAFETWIGRXGEXPVPLTLMSLSV--S 71
      QY 48 EFLVGLLIGALLASLVGFLIDCYGRKQALIGSNVLLAGSLTGL--AGSLAMVLG 104
      DB 72 IFAVG-MIGSLVLYXIGNL---GRKXAMLVNNVLAAGGLMLGKAXXSFEMLLIG 126
      QY 105 RAVGFAISLSMACCIYSELVGPORGVLVSLYEAGITVYGIISVALN-ALAGTPMG 163
      DB 127 RFIIGLYCGLSSGVPMVYGEISPTALRGALGTINOLGIYIGILIAOVGLDLSLGNESL 186
      QY 164 WRHMGATAPAVLQSLSLFLPAG-----TDETATRKDLIPLOG--G 204
      DB 187 WPLLGLGVLPALDILLPCBPESPRYLLINNEARAKKALQRLGTADVSGEVAEMK 246
      QY 205 EAPKLPGRPRYSFLDFRARDNKRRTTVGLVLFQOLTGOPNVLYCASTIFFSSVGFH 264
      DB 247 DESRMXSEKXSVLELFRSR-XRQPVIIAIVLQUSQLGNAVYVSISIEKAGV- 304
      QY 265 GGSASVLAISVGLGAVKVAATLTAMGLVDRAGRALL-----AGCALMALSVSGIGLV 318
      DB 305 --GQPVATIGAGVNVTVETVVSFVERAGRTTLLILGAGMACAV-----LMT 353
      QY 319 FAVPMDSGSCSLAVPMAGOTGLPDSGLDSSLPPIPTNEDQREPLISTAKTKRPHR 378
      DB 354 LA-----355
      QY 379 RSGDPSAPRLALSALPGLPARGHALLRMTALLCLMVFVSASFSGFPYTWLVSEI 438
      DB 356 -----LALDQV-----WMSYVSIVAITGFAVFEVGPPIPMFIVAEL 395
      QY 439 YPEIRGRAPFAFCNSFNMANLFISFLDILGTSMTFLVGLTAVLGIFLYLFVP 498
      DB 396 FSGCPREAAIVAGFSNMTSNFTVGLLFQYIAELIG-PVFIYFVALLLFIETFLKVP 454
      QY 499 ETQGOSLAETDOOROK 514
      DB 455 ETGKRTFDEIAAFRK 470
  
```

RESULT 3
 US-09-299-549-10
 Sequence 10, Application US/09299549
 Patent No. 6136347
 GENERAL INFORMATION:
 APPLICANT: Tartaglia, Louis A.
 APPLICANT: Meng, Xun
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 TITLE OF INVENTION: GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkiohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-299-549-10

Query Match 15.4%; Score 427; DB 4; Length 493;
Best Local Similarity 25.9%; Pred. No. 3,7e-30;
Matches 144; Conservative 80; Mismatches 184; Indels 148; Gaps 15;

QY 8 LPLCASVSLGGLTFGEYELAVISG-----ALLPDLQDFGLSCLEQ 47
DB 14 LVLAVLIALGSLFOGYVNLINAPQKVIENAFYETWLGKXGKSPVTLTLMSSLSV--S 71
QY 48 EPLVGSLLGLALLSLVGGFLIDCYGRKQAIIGSLNVLVLAGSLTGL--AGSLAMVLV 104
DB 72 IFAVGG-MIGSFLVXIGNRL---GRKXAMLVNNVLAIGLMLGAKKXSSPEMLILG 126
QY 105 RAVVGFATISLSMACCIYVSELVGPVRGVLSLYEAGIVGLISALNV-ALAGTPMG 163
DB 127 RFIIGLYCGLSSGVVPMYVGEISPTALRGALGTNLQIGIVIGILIAQVGLDLSLGNESL 186
QY 164 WRHFGWATAPAVLQSLSLFLPAG-----TDETATHKDLPILOG-----G 204
DB 187 WPLLLGLTGVALLQLLLPFCPEPRYLLINKNEGRARAKALQRLKGTADVSGEVAEMK 246
QY 205 EAPRLGCRPRYSFLDLFRANDMRGRTVGLGLVLPQULTGPPNVLCTASTIFSSVGFH 264
DB 247 DESRXMKSEKXSVLELFRSR-XYRQVITAIYVLSQQLSGINAVFYSTISFEKAGV- 304
QY 265 GGSAAVLASVGLGAVKVAATLTAMGLVDRAGRALL-----AGCALMALSVSGIGLV 318
DB 305 --GQPVYATIGAGVNTVFTVSVFVERAGRRTLLHLLGGMAGCAV-----LMT 353
QY 319 FAVPMDSPLSLAVPMNATGQGTGLPDGSLQDSSLPPIPTNEDQREPISTAKTKRHP 378
DB 354 IA----- 355
QY 379 RSGDPSAPRLALSSALPGPPLPARGHALLRMTALLCLMVFVSASFSGPVTWLVLSI 438
DB 356 -----LALLDQVP-----WMSYIVAIIFGFVAEFEGPGPIPMFTIAEL 395
QY 439 YPVEIRGRAPAFCSNFNMAANLFISLFLDLTIGTIGLSTWFLYGLTAVLGLGFIYLFV 498
DB 396 FSGQPRPAALIVAGRSNMTSNITVGLLFFQYIAELLG-PYVFYIYFAVILLLFFITFLKVP 454
QY 499 ETKGQSLAEIDQGFQK 514
DB 455 ETKGRTFDEIAAARFK 470

RESULT 4
US-09-031-392-6

Sequence 6, Application US/09031392
Patent No. 5942398
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkiohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-031-392-6

Query Match 14.7%; Score 407.5; DB 2; Length 509;
Best Local Similarity 25.5%; Pred. No. 2,1e-28;
Matches 143; Conservative 81; Mismatches 196; Indels 141; Gaps 15;

QY 5 P-----VPLPCASVSLGGLTFGEYELAVISGALLPLQDFGLSCLEQ----- 48
DB 15 PPGQRTGTLVAVESAVLSLQFQYINAPQKVIENAFYETWLGKXGKSPVTLTLMSSLSV--S 71
QY 49 -----FLVGSLLGLALLSLVGGFLIDCYGRKQAIIGSLNVLVLAGSLTGLAGSLA- 99
DB 75 GTLTLMALSAIFSVGGMISFLIGIISOMLGRKRAMLVNNVLAIGLMLGAKKXSSPEMLILG 126
QY 100 --WLVGAAVAVFAISLSMACCIYVSELVGPVRGVLSLYEAGIVGLISALNV-A 156
DB 135 YEMILIGRFLIGAVSGTLVPMYGEIAPTHLRGLGTNLQAIYIGILIAQVGLGIES 194
QY 157 LAGTPMGWRHFGWATAPAVLQSLSLFLPAGTD-----ETATHKDLPILOG----- 203
DB 195 LIGTASLWPLLLGLTGVALLQLLLPFCPEPRYLLINKNEGRARAKALQRLKGTADVSGEVAEMK 246
QY 204 -----GEAPRLGCRPRYSFLDLFRANDMRGRTVGLGLVLPQULTGPPNVLCTAST 264
DB 255 GVLAEKKEKKERERER-LSLQLLGSRTN-RQPLIIAVVLSQQLSGINAVFYSTISFEKAGV- 304
QY 257 IFSSVGRFGSSAAVLASVGLGAVKVAATLTAMGLVDRAGRALLLACALMALSVSGIGL 316
DB 313 IFETAGV--GQPVYATIGAGVNTVFTVSVLVERAGRRTLLHLLGGMAGCAV-----LMT 353
QY 317 VSFVPMDSPLSLAVPMNATGQGTGLPDGSLQDSSLPPIPTNEDQREPISTAKTKRHP 378
DB 361 CGCAILMTVALLL-----ERVPMAS----- 381
QY 377 HPSGDPSPAPRLALSSALPGPPLPARGHALLRMTALLCLMVFVSASFSGPVTWLVLS 436

Db 382 -----YVSIAIFGFVAFEEIGPPIPWFIYA 408
 QY 437 EIYPVEIRGRFAFCNSFPMANLFLISFLDLIGTIGLSTFLLYGLTAVLGIGFI--- 493
 Db 409 ELFSQGRPRAMAVAGFSNMTSNFTIGMGFYVAEAMG-PYVFLIF---AVLLGFIET 464
 QY 494 YLFVPEETKQSLAEIDDOFOR 514
 Db 465 FLRVPEETRGRTFDOISAAFR 485

RESULT 5

US-09-299-549-6
 ; Sequence 6, Application US/09299549
 ; Patent No. 6136547

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 TITLE OF INVENTION: GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESS: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,549
 FILING DATE: 26-APR-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkielejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELE: 200154

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 509 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-299-549-6

Query Match

Best Local Similarity 14.7%; Score 407.5; DB 4; Length 509;
 Matches 143; Conservative 81; Mismatches 196; Indels 141; Gaps 15;

QY 5 P-----VLPCLASVSLIGCTFYELAVISGALPLQDLDFGLSCLDE----- 48
 Db 15 PPOORVGTGLVAVFSALVGLGFGYNIGVINAPKVIQOSVNETMIGRGPGPSSIPP 74
 QY 49 -----FLVGSLLGLASLVGFLIDCGKQOAILGSNVLVLGSLTGLAGSLA- 99
 Db 75 GTLTTLALSAIRSVGGMISSEFLIGITISOWLGRKRAMLVNNVLAVLGSLGSLA- 134
 QY 100 --WLVGRAVAFATISLMAACIYSELVGPORGVLSLEAGITVGLISVALNY-A 156
 Db 135 YEMILICRFLIGAVSGITSLVPMYVGEIAPTHLKGALGTINOLAIVIGILTAOVIGLES 194

QY 157 LAGTPMGWRHMGWATAPAVLOSLSLFLPAGND-----ETATHKDLIPLOG----- 203
 Db 195 LLGTASLWPLLLGLTGLVLRALLQVLPLPCPESPRIYITQNEGPARKSLKRLTGWADVS 254
 QY 204 -----GEARKLGRRRYSFLDLFRARDNMRGRTVGLVLFQOLNGQPVNLGAST 256
 Db 255 GVLAELDKERKLERERD-LSLLOLGSRTH-RQGLIAIVQLSOLSGINAVNYITS 312
 QY 257 IFSSVGFHGGSSAVLASVGLGAVYAATLTAMGLVDRAGRALLAGCALMASVSGIGL 316
 Db 313 IFETAGY---GQPARATIGAGVNTVFLLVSLLVERAGRRLTLGLA-----GM 360
 QY 317 VSRVAVPDSDSCIAVPMATGOTGLPGDSGLDQSLPPIPTNDQREPLISTAKKTP 376
 Db 361 CGCATLMTVALLL-----ERVAMS----- 381
 QY 377 HPRSGDPSAPRRLALSALGPPPLPARGHALLRMTALLCLMVVSASFSGPVTWLVLS 436
 Db 382 -----YVSIAIFGFVAFEEIGPPIPWFIYA 408
 QY 437 EIYPVEIRGRFAFCNSFPMANLFLISFLDLIGTIGLSTFLLYGLTAVLGIGFI--- 493
 Db 409 ELFSQGRPRAMAVAGFSNMTSNFTIGMGFYVAEAMG-PYVFLIF---AVLLGFIET 464
 QY 494 YLFVPEETKQSLAEIDDOFOR 514
 Db 465 FLRVPEETRGRTFDOISAAFR 485

RESULT 6

US-09-031-392-5
 ; Sequence 5, Application US/09031392
 ; Patent No. 5942398

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESS: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Melkielejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/072001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELE: 200154

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 494 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-031-392-5

Query Match 14.6%; Score 404.5; DB 2; Length 494;
 Best Local Similarity 23.9%; Pred. No. 3.8e-28;

ADDRESSER: No. 59587270 No. 59587271disk of No. 5958727th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,692
 FILING DATE: 12-SEPT-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4944,200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-867-9655
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 524 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 5958727e
 US-08-928-692-12

Query Match 14.3%; Score 394.5; DB 2; Length 524;
 Best local similarity 23.6%; Pred. No. 3.3e-27;
 Matches 138; Conservative 83; Mismatches 182; Indels 181; Gaps 16;

QY 15 SLUGLFGELAVISGA-----LLEPLD-----FGLS- 43
 DB 17 AVLSFGDGVGINAPQOVIISHYRVLGVPLDDRAKANNVINSTDELPTISTSNP 76
 QY 44 -----CLEOEL-----VSLGALLALIASLVGFLIDCYGRKQALIGSNVL 86
 DB 77 KETPAEETVAAAOILTMLMSLVSSFAVGCMATSPFGMLGDTGRKMLVANIISL 136
 QY 87 AGSLTGL--LAGSLAMLVGRAVGFALISSMACCIYSELVGPORGVLVSLYEAGI 143
 DB 137 VCALLMGFSKLGPSHLIIAGRSISGLYGLISGLVPMYIGELAPTLRGLGTFHQLAI 196
 QY 144 TVGILISTAL-----NYALACTPMGWRHMFGMATAPAVLOSILFLPAG----- 188
 DB 197 VTGILISOIIGLEFILGNVDL-----WHILGLSGVRAIIQSLLFPCEPSPRYLTKL 250
 QY 189 TDEATHDILPLOG-----GEAPKLGPRGRYSFLDLFRARDNMRGRTVGLG 237
 DB 251 DEEKAKOSLKRIRGYDVTXKIDINEKREREREASSQKVSIIQLF-TNSYSROILVALM 309
 QY 238 LVLFQOLGQPNVLCYASTIFSSVGHGSSSAVLAVGCAVVAATLTAMGLVDRAGR 297
 DB 310 LHVADQPSGICINGIFYSITIFOTAGI--SKPYATIGVAVVMVFAVSFLVEKAGRR 366
 QY 298 ALLIAG-----CALMALSVSGIGLVSAFVPMDSGSCLAVPATOTGTGPGSGLIQDS 351
 DB 367 SLFLIGMSGFVCAIF-----MSVGLV-----LNKLF 393
 QY 352 SLPPPTNEDQREPIISTAKKTRKPHRSGDPSAPPRLALSSALPGPPLPARGHALLRMT 411
 DB 394 S-----MWSY 399
 QY 412 ALLCLWVSASFSGPGVTVLVSEIYVPEIRGRAPFCNSFNMAANDFISLFLDLIG 471
 DB 400 SMATILFVSFFETGPETPMFVAEEFFSGPRAALAIARSNWTCTIVALCFOYIAD 459
 QY 472 TIGLSWTFLLYGLTAVLGLGFIYLVFETPKGSLAIEDOQFOK 515

DB 460 FCGPYVFFLFAVGLAFTL-FTFEKVPETKGSFEELIAEFOK 502

RESULT 9
 US-09-031-392-4
 Sequence 4, Application US/09031392
 Patent No. 5942398

GENERAL INFORMATION:

APPLICANT: Tartaglia, Louis A.

APPLICANT: Meng, Xun

TITLE OF INVENTION: NUCLEIC ACID MOLECULES

NUMBER OF INVENTION: ENCODING GLUTEX AND USES THEREOF

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,392

FILING DATE: 26-FEB-1998

ATTORNEY/AGENT INFORMATION:

NAME: Meikiejohn, Ph.D., Anita L.

REGISTRATION NUMBER: 35,283

REFERENCE/DOCKET NUMBER: 07334/072001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 534 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-031-392-4

Query Match 14.1%; Score 389.5; DB 2; Length 534;
 Best local similarity 24.3%; Pred. No. 9.4e-27;
 Matches 146; Conservative 78; Mismatches 194; Indels 183; Gaps 16;

QY 3 HSPVPLCLASVSLGGLTGYELAVIS----- 30
 DB 12 HLTGLVLSVFTAVLGFQYGSIGVINAPOKVIHAHGRMLGATPMVNRHATNTSRDNT 71
 QY 31 -----GALLP-----LQDFGLSCLDEFLVGLSLLGALLAS 62
 DB 72 TVTPTITGTRMWSSEGLTASAGFEDPTVSPHILTWMSLS-----VSMFAVGAMVSS 124
 QY 63 LVGGLFLIDCYGRKQALIGSNVLVLAGSLTGLA-----GSLAMLVGRAVGFALISSMAC 119
 DB 125 FTVGMIGDRLGRKAMLVVNVLSIAGNLMGLAKMGPSHIIINGRAITGGLYCGISGLV 184
 QY 120 CIYSELVGPORGVLVSLYEAGITVIGILSYALNT-ALAGTPMGWRHMFGMATAPAVLO 178
 DB 185 PMYVEVSPTALRGALGTIHLQALAVTGILISQVLGDLFLGNDLMLPLGLSGVALLO 244
 QY 179 SLSLFLP-----ACTDETAHKKDLIPILOGEAP-----KLGPGRPRYSF 218
 DB 245 FFLLELCESPRYLYIKLKVEEA-KSLKRLRGNCDDPKMEIAMEKEKDAASEKRVSI 303
 QY 219 LDLFRARDNMRGRTVVGGLVLFQOLTGQPNVLCYASTIFSSVGHGSSSAVLAVSGLA 278
 DB 304 GOLF-SSSKRYQAVIVALLVVOISQDFSGINATFYSTINIFORAGV--GQPVYVATIGV 360

```

OY 279 VVVAATLTMTAGLVDRAGRALLLAGCALMALSVSGIGLVSPFVPDPDGGSPCLAVPATGQ 338
      | : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 361 VNTVTVTVISVFLVERKKGRSLFLAG--LMCHLIS-----AVANTVG----- 399
OY 339 TGLPGDGLLDSSLPPIPTNEDQREPIITSTAKTKRPHRSGDSPAPRLALSALPCP 398
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 400 -----LVLLSQP----- 406
OY 399 PLPARCHALLRMTALCLWVFAFSFGCPVTMLVLSSEIYPVEIIGRAFAFCNSFNMAA 458
      | : : : | : : | : : | : : | : : : | : : : | : : | : : | : : |
Db 407 -----AMMSYVSMVAIFLFVILFEVEGPGCIIPMFIVALELSPGGRPAIIAAGFCNMAC 459
OY 459 NLFISLSF---LDLIGTIGLSMTFLLYGLTNVGLGFIYLLFVPETKGQSLAEIDQFOKR 515
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 460 NFIYVCMCFIADLCG---PAYFVVVFAVALLLVFELFAYLVAKPKETKGSFEETIAAFRRK 515
OY 516 R 516
Db 516 K 516

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Db	12	HLTGTLVSVFTAVLGGFFQGYCSGLGVINAPDKVIEAHNGRMIGALPMVRIHAINTSRDNT	71
Qy	31	-----GALLP-----LQDLGSLCSLEDFYVGLSLILGALLAS	62
Db	72	ITVTITPGTEAMSGSSEGTFLAPASAGFEDPTSPHILITMWSL-----VSMRVAQGMSS	124
Qy	63	LYVGGLIDCYGRKQATILGSLVLLGASLTGLA---GSLMWLVGRVAVGFAISLSMNC	119
Db	125	PLTVGWIGDRIGRKAMLVVNVLSIAGLLKGLAMGCPSHIILIACRAITGLCYGLSSGLV	184
Qy	120	CIYSELYGPRQKGVLSLYEAGITVGLIILSTYALNT-ALACTPMQWRHMFQATAPAVLQ	178
Db	165	PMYSEVSPITALRGALGTLHQIALIVTICILISQVIGLDFLLGNDLWPLLLGLSGVALLQ	244
Qy	179	SLSLFLP-----AGTDEFATHKDLIPLOGGEAP-----KLGPRPRYSF	218
Db	245	PFLLTLCPESPRLYIKLGVEEA-KKSRLRRCNCPMKKEIAEMEKKEQEAASEKRVSI	303
Qy	219	LDLFRANDMRGTTTGLGLVLRQOLLGQPRVLCYASTIFPSVGRHGSSAVLASVGLA	278
Db	304	GOLF-SSSKYRQAVYVLAALWQISOFSGINAIFYSTNIIPRAGY--COPYVYATITGVV	360
Qy	279	VVVAATLTLMGLDRAGRRALLLAGCALMALSVSGICLVSEAVPMDSGPSCLAVPNATQ	338
Db	361	VMTVFVIVSVFLVEKKGRRSLPLAG-LMGHLIS-----AVANTVG-----	399
Qy	339	TGLPQDSGLQDSSLPIPIPTNEDQREPIILSTAKKTRHPNSGDPASAPRRLLSSALPCP	398
Db	400	-----LVLLOSF-----	406
Qy	399	PLPARGHALLRMTALLCLMVFASFSFGCPVTMLVLSIYPVIEIRGRAFAFCNSEMMA	458
Db	407	-----AMMSYVSMALFLFVIFFEVGPQGIIPMFIVALELSSQGRPAALAAAGPCNMNC	459
Qy	459	NLFISLSF---LDLIGTIGLSWTFLLYGLTAVLGLGFIYLFVPEPKGOSLAEIDQFOKR	515
Db	460	NFTVGCMFQIADLCG---PYVAVVVAVLLLVLFLLFVLYKVPETKGSFEEIAAFAFRK	515
Qy	516	R 516	
Db	516	K 516	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

NAME: Tang, Henry Y.S.
 REGISTRATION NUMBER: 29,705
 REFERENCE/DOCKET NUMBER: A29927-50/29910
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2586
 TELEFAX: 212-765-2519
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 492 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Human
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 1..492
 OTHER INFORMATION: Facilitative glucose transporter
 US-08-355-844-3

Query Match 13.2%; Score 364.5; DB 2; Length 492;
 Best Local Similarity 24.6%; Pred. No. 14e-24;
 Matches 136; Conservative 82; Mismatches 194; Indels 141; Gaps 16;

8 LPLCAVSLGLGTFEYELAVISG-----ALPLQD--FGISGLE 46
 12 LMLAVGAVGLSLQFGYNTGVINAPQKIEEFYNTQWHRGESILPTLTTLTMSLS--- 68
 47 QEFVLSLLGALLASLVGGLFIDCYGRKQALGSLNVLASLTG--LAGSLAMVL 103
 69 ----VAIFSVGMIGSFSVGLFVNFRGRNSMLMNLAFVASAVLMGFSKLGSEFEMIL 124
 104 GRAVGFALISLMACCIYSELVGRORGVLSYEAGITVIGLSTALNY-ALAGTPW 162
 125 GRFIIYVCGITGTGFVPMYGEVSPFAFGALGTLHQLGIVVGLILAOVFGLDISGNKD 184
 163 GWRHMGWATAPAVLOSILFLPAG-----TDETATKDLIPLOG----- 203
 185 LMPLLSLITFIPALLOCIIVIPCPESPRLINRNEENRAKSVLKKLGTADVTHTDLEM 244
 204 GEAPKLGPRPRYSFLDFRARDNMGRITVIGLVLFQOLTGDPNVLCTVSTIFSSVGF 263
 245 KEESROMREKKVTILFLRS-PAYROPILIAVVLQSLQSLGNAVYYSSTISFEKAGV 303
 264 HGGSSAVILASVGLGAVKVAATLTAMGLVDGRALLAGCALMALSVSGIGLSEFAVPM 323
 304 ---QCPVYATIGSGIVNTAFTVVSFLFVERAGRTLHLIGLAGNA----- 345
 324 DSGPCLAVPNATGGVGLGDSGLDSSLPLPIRINEDOREPILSTAKTKPPHPRGDP 383
 346 -----GQA-----ILMTI----- 353
 384 SAPRLIALSSALPPPLPARGHALLRMTALLCLMVFVASFGPPTWVLSTIYEVI 443
 354 ----ALALIELQF-----WMSYLSIVAIFFGVAFEEVGPPIPMWIVALESGGP 399
 444 RGRAFAFNSFMANLFLISFLDLIGTIGLSWTEPLVGLTAVLGIGFT--YLIVPETK 501
 400 RPAIIAIVAGFSNMTSNFTYVCMFOYVEQLCG-PYVFILF--IVLIVLFFIRTYFKVPETK 456
 502 GQSLAEIDQPOPK 514
 457 GRTDEIATASGRQ 469

RESULT 12
 PCT-US95-16126-3
 ; Sequence 3, Application PC/TUS9516126
 ; GENERAL INFORMATION:
 ; APPLICANT: Fischberg, Jorge

APPLICANT: Czegledy, Ferenc
 APPLICANT: Iserovich, Pavel
 APPLICANT: Li, Jun
 APPLICANT: Cheung, Min
 TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
 TITLE OF INVENTION: STRUCTURE
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
 STREET: 30 Rockefeller Plaza
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10112-0228
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/16126
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/355,844
 FILING DATE: 14-DEC-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Tang, Henry Y.S.
 REGISTRATION NUMBER: 29,705
 REFERENCE/DOCKET NUMBER: A29927-50/29910
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-408-2586
 TELEFAX: 212-765-2519
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 492 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Human
 FEATURE:
 NAME/KEY: Peptide
 LOCATION: 1..492
 OTHER INFORMATION: Facilitative glucose transporter
 US-09-652-292-2

Query Match 13.2%; Score 364.5; DB 5; Length 492;
 Best Local Similarity 24.6%; Pred. No. 14e-24;
 Matches 136; Conservative 82; Mismatches 194; Indels 141; Gaps 16;

8 LPLCAVSLGLGTFEYELAVISG-----ALPLQD--FGISGLE 46
 12 LMLAVGAVGLSLQFGYNTGVINAPQKIEEFYNTQWHRGESILPTLTTLTMSLS--- 68
 47 QEFVLSLLGALLASLVGGLFIDCYGRKQALGSLNVLASLTG--LAGSLAMVL 103
 69 ----VAIFSVGMIGSFSVGLFVNFRGRNSMLMNLAFVASAVLMGFSKLGSEFEMIL 124
 104 GRAVGFALISLMACCIYSELVGRORGVLSYEAGITVIGLSTALNY-ALAGTPW 162
 125 GRFIIYVCGITGTGFVPMYGEVSPFAFGALGTLHQLGIVVGLILAOVFGLDISGNKD 184
 163 GWRHMGWATAPAVLOSILFLPAG-----TDETATKDLIPLOG----- 203
 185 LMPLLSLITFIPALLOCIIVIPCPESPRLINRNEENRAKSVLKKLGTADVTHTDLEM 244
 204 GEAPKLGPRPRYSFLDFRARDNMGRITVIGLVLFQOLTGDPNVLCTVSTIFSSVGF 263
 245 KEESROMREKKVTILFLRS-PAYROPILIAVVLQSLQSLGNAVYYSSTISFEKAGV 303

[illegible]

```

Db      75 VSNFPFGGFIGSLLVGLVYKFGKQKALLFNFNISIVPALLMGCSRVATSPFELLIIISRL 134
QY      108 VGFALISLSRACCIYVSELVGPORGVLSLYEAGITVGILIS--YALNTALAGTPWGR 165
Db      135 VGIACAVSSNVVMYIGELAPKNLRGALGVVPOLFTVGIIVAOIFGLNLLANDV-GMP 193
QY      166 HMGMTATAPAVLOSLSLFLPAG-----TDEATHKDLIPLOGGEAPKLGPRPR-- 215
Db      194 ILGLGVPAIDOLLPLFPESPRLYLIOKKDEAANKALQITLGWDSVDREVAIRDE 253
QY      216 -----YSFLDLFRARDNRGRFTVGILVLYEQLTGQPNVLCYASTIFSSVFGHG 266
Db      254 DEAKAAGFISVYKLFPMR-SLRQQLSIYILMGQQLSGVNAIYYAAQIYLSAGV-PE 311
QY      267 SSAVLASVGLGAVKVAATLTAMGLVDRAGRALLAGCALMALSVSGIGLVSPAVPMDSG 326
Db      312 EHVOYTAGTGAANNVWTPCAVFEVVELLGRLLLLGFSICLI----- 354
QY      327 PSCLAVPNMTGOGVGLPDSGLDQSSLPPIRTNDEQREPLITSTAKTKPHRSGDPSAP 386
Db      355 -ACC-----VLTAA----- 362
QY      387 PRLATSSALPGPELPARGHALLRMTALLCLAVVSAFSGEPVWLVSEIYVEIRGR 446
Db      363 --LAIDDTYSMP-----YISIVCSYIYGHALGSPRIALLIT-IFLOSRRS 409
QY      447 AFAFCNSFMNAANLFLISLFDLIGTIGLSMTFLLYGLTAVLGLGFIYLFVPETKQSLA 506
Db      410 AEWGSGVHMLSNFTVGLLFPFIQEGIG-PYSFIVEAVICLITITIIYIFLIVETKAKTFI 468
QY      507 EIDDOFOR 514.
Db      469 EIINQIFTK 476

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 500 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-299-549-7

Query Match 12.6%; Score 349.5; DB 4; Length 500;
 Best Local Similarity 25.0%; Pred. No. 3.2e-23;
 Matches 137; Conservative 85; Mismatches 199; Indels 127; Gaps 17;

QY 7 VLPICASVSLG-GLTGYELAVI-SCALLPLQI-----DEGLSCLEDEFL 50
 Db 16 VLALFTLIAFSSFOYGVNAVNSPALMOOFYNETYYGRTEFMEDEFLTL-MSVT 74
 QY 51 VSLGLLALSLVGFLLDCYGRKQALIGSNVLLAGSLTIG---LAGSLAVLYGRAV 107
 Db 75 VSMFEGFGLSLVGLPLNKFGRKALLFNFIPTVPALMGCSRYATSEFLIIRSL 134
 QY 108 VGFALISSMACCIYSELVSPRGVLSVLEAGITVGLLS--YALNVALAGTPGWR 165
 Db 135 VGICAGSSNVVPMYLGELAPKRLGALGVLPQLFITVGLVNOITGLRNLANVD-GWP 193
 QY 166 HMFQMATPAVLOSLSLFLPAG-----TDEPITHDLPLOGGEAPKLGPRPR- 215
 Db 194 ILGLTGVPAALQILLFFPESPRLYLQKKDEAAKALQTLRGMDSDVDEVAIRQE 253
 QY 216 -----VSFLDFRADNMRGRTVGLVLPQOLTGQPNVLCYATSESSVGRHG 266
 Db 254 DEAKAGFISVLFKFRMR-SLRQQLSIYVLMGQQLSGVNAIYYADQIYLSAGV-PE 311
 QY 267 SSAVLASVGLGAVKVAATLTAMGLVDRAGRALLAGCALMALSVSGIGLVSEFVPMDSG 326
 Db 312 EHQYVYTAGGAVNVMTFCVAVFVELIGRLLILGFSICLI----- 354
 QY 327 PSLAVPNATGQGLPDSGLLODSSLPPIRTNEDOREPIITSTAKTKPHRSGDPSAP 386
 Db 355 -ACC-----VLTAA----- 362
 QY 387 PRALSSALPPLPARGHALLRWLALCLMVFVSAFSGFGPVALVLSIYPIEIRGR 446
 Db 363 --LALDVTWMP-----YISIVCVSYIGHALGSPHALIT-IFLOSSRPS 409
 QY 447 AFAFCSPFMAANLFTSLFLDILGIGLSTWTFLLYGLTAVLGLGYLFPVETKGSIA 506
 Db 410 AMWGSVWMLSNPTVGLPFTQEGIG-PYSFVFAVICLITITIVFLYVETKAKTFI 468
 QY 507 EIDQOFK 514
 Db 469 EINDIFTK 476

RESULT 15
 US-08-928-692-10

Sequence 10, Application US/08928692
 Patent No. 5958727

GENERAL INFORMATION:

APPLICANT: Brody, Howard

APPLICANT: Yavel, Deborah S.

APPLICANT: Hansen, Kim

TITLE OF INVENTION: Methods for Modifying the Production of

NUMBER OF SEQUENCES: 80

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 59587270 No. 5958727disk of No. 5958727ch America, Inc.

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,692
 FILING DATE: 12-SEPT-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4944,200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 488 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-692-10

Query Match 12.0%; Score 331; DB 2; Length 488;
 Best Local Similarity 23.1%; Pred. No. 1.4e-21;
 Matches 130; Conservative 84; Mismatches 190; Indels 158; Gaps 15;

QY 6 PVLPLCASVSLGTLTGYELAVISGALLPLDID-----FGL-SCL-----EDEFV 51
 Db 13 PVLIVYVTLTGLPQFQGLHAEINAPQAVITCERKSIHSTTRGLPQCIIPMPSPQGLV 72
 QY 52 GSL-LGALLASLVGCPFLDCYGRKQALIGSNVLLAGSLTIGLASSLAVLYGRAVGF 110
 Db 73 SSITLGLLALLAGPSTKGRIFLTATITIFLIGPIAETFPASIPVLSMGRILSGV 132
 QY 111 AISLSSMACCIYSELVSPRGVLSVLEAGITVGLLSYALNVALAGTPGWRHMFQ 170
 Db 133 GAGASIVGPIYISEIAPSAKGLGCAFTQIMTNVGLITGSLGFLSGM-WRVILAI 191
 QY 171 ATAPAVLOSLSLFLPAGTDEPITHK----- 197
 Db 192 AGAIGCELGLFLVPSPIVLADHOKGNVAROYLQIRGRDADIEPEVGRMISAPEH 251
 QY 198 -----LPLLOGGEAPKLGPRPRYSFLDFRADNMRGRTVGLVLPQOLTGQPN 249
 Db 252 SSSEGLSLSPSGNMPKCPPTMMRAITDSF-----YRPAITAVGVMSQOFTGVNS 306
 QY 250 VLCYASTIFSSVGRHGSSAVIASVGLGAVKVAATLTAMGLVDRAGRALLAGCALMAL 309
 Db 307 IIMVSVSLQTI--LPTTAAALSVIISAINLVITLACSPLPDKIGRSCLL-----L 356
 QY 310 SVSGIGLVSEFVPMDSGSPCLAVPNATGQGLPDSGLLODSSLPPIRTNEDOREPIIS 369
 Db 357 SISGMGLNSVLL-----ALAI----- 372
 QY 370 TAKTKPHRSGDPSAPPRIALSSALPPLPARGHALLRWLALCLMVFVSAFSGFGP 429
 Db 373 -----YNLKAISIAIVLLEFVASFPAGLGP 397
 QY 430 VTMLVLSIYPIEIRARAFCSFMAANLFTSLSF--IDLIGTIG-LSWTFLLGLT 485
 Db 398 VFTILASELVGEAVGAAQSWALGANMIATFVIAQFPPMLNDLGGSGKIYWIFA--AMA 455
 QY 486 AVLGLGYLFPVETKGSIAE 507
 Db 456 CLG-SPIYMWVETKGRANAD 476

Search completed: March 15, 2002, 07:03:47
 Job time: 4438 sec

THIS PAGE BLANK (USPTO)